Potential of Imprecision: Exploring Vague Language in Agent Instructors

Leigh Clark*, Khaled Bachour*, Abdulmalik Ofemile†, Svenja Adolphs†, Tom Rodden*
*Mixed Reality Lab †School of English
University of Nottingham, UK
{psxlc, khaled.bachour, aexacof, svenja.adolphs, tom.rodden}@nottingham.ac.uk

ABSTRACT
As we find greater potential for agent instructors, we must be aware of how their language use can affect the user and interaction as a whole. This study investigates the use of intentionally imprecise or *vague* language as a communicative strategy to mitigate the impact of instructions. We look at the effects it has on improving the perception of agents and user performance. A series of assembly tasks were ran in which users constructed Lego models with the spoken instructions of vague and non-vague agents. Results show that though the non-vague agent was seen as more direct and authoritative, responses to other attributes and performance were much more varied. Findings suggest there is potential for vague language human-agent interaction, though there are several obstacles in agent design to overcome first.

Author Keywords
Human-agent interaction; instructions; vague language; communication strategies

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Interaction Styles

INTRODUCTION
The prevalence of agents in our lives continues to rise and our interactions with them are becoming more complex on both a communicative and social level. Conversational and relational agents in particular aim to achieve a sense of rapport with their users [3, 4]. These represent a move into the emerging science of human-agent collectives (HACS) and with them present new challenges as to how agents best convey information and an awareness of how their use of language can affect the interaction as a whole. In HACs humans and agents can take on a range of varying roles, and successful communication becomes crucial to the effective operation of the collective.

HAI '14, October 29 - 31 2014, Tsukuba, Japan.

http://dx.doi.org/10.1145/2658861.2658895

This study takes a similar but new approach with one particular arrangement of roles: it focuses on agents giving direct and vague verbal instructions to human participants. It also hopes to assess the effect of these on participants when agents employ politeness strategies and some issues that arise from the use of natural language to achieve this communication.

BACKGROUND

Agents as Instructors
When talking about Human-Agent Collectives, successful communication requires that humans be open to being directed [22], and able to engage with agents at a peer level [17]. Agents are capable of dealing with some types of information in quantities and complexities that would overwhelm humans [2], and it is in these situations that they are ideally suited for a role as instructor, making quick decisions with vast amounts of data. A lot of work has been done on the role agents can play in the management of complex and information rich situations such as emergencies [21] and damage control [6]. Agents have also been shown to be able to hold more advisory roles such as a personal tutor [14] or by assisting patients and medical staff in diagnoses [8, 11].

Vague Language
While machine communication is by design direct and unambiguous, everyday human communication often contains varying degrees of uncertainty know as vague language (henceforth VL) [9, 10]. This can arise from genuine uncertainty but when used deliberately is a communicative strategy used to achieve functional and relational goals simultaneously. For example, a student answering a mathematics questions in classroom may respond with, "but it's around 50 basically?" [20]. Here the speaker conducts the functional goal of answering a question given by a teacher, while also fulfilling the relational goal of protecting oneself from full commitment to the answer and potential error by being imprecise. VL is seen a wide array of other contexts such as medical examinations [1], academic conferences [25] and the workplace [15].

The example above shows the speaker using VL as a politeness strategy [5]. Politeness strategies allow speakers to convey information such as instructions and requests without encroaching upon the listener's independence and freedom of action without impediment known as *face* [12, 13]. Instructions in particular create an imbalance of power that can potentially create a social gulf between two speakers, so polite communication can be used to convey instructions through the medium of imprecise language.

There has been some work on incorporating politeness strategies into HAI. When used in the classroom polite agents were seen to improve learning outcomes [26]. Similarly, when used in advice giving robots for a baking task it was shown to make them appear more likeable, considerate and less controlling [24]. This explored the successful use of hedges and discourse markers as polite communication. Hedges such as *kind of* and *sort of* allow a speaker to express uncertainty and avoid being assertive [16, 19]. Discourse markers such as *basically* and *like* are able to soften commands and distance the speaker from the information they are delivering to the listener. Though not described as such, these both represent features of VL through their deliberate imprecision with a purpose of both interactional and relational success. For the purposes of this paper we adopt a similar approach in the creation of a linguistic framework. As both hedges and discourse markers were used for the same purpose with similar success we combine the two under the banner of VL. This also allows us to include words that may usually be assigned to other categories such as fillers in our framework, so long as it is used in accordance with VL definitions.

While the use of VL was able to make a difference perceived attributes of robots in advice giving, it is unknown whether the same can be achieved with instruction giving agents. Instructions represent a more rigid information structure in which there is a closed set of outcomes, particularly those in assembly where each step is dependent on the ones preceding it.

In investigating VL use with agent instructors we devised four hypotheses based on previous literature. Firstly, we envisage human users will rate a vague agent as more likeable, friendly, trustworthy and sociable than a non-vague agent (H1). Similarly, we believe the non-vague agent will be rated as more controlling, authoritative, clear and direct (H2). It is thought the protection of the user's face and attempts to create a socially level discourse [7] will create a notable difference in how the different agents are perceived. We also predict that user performance will increase when following the vague agent instructions by creating an equal relationship [18] in a similar vein to what they expect from human instructors (H3). Finally, we predict that the introduction of an external stress factor will see a reduction in the differences seen in H1-H3 (H4).

METHOD

To test the hypotheses we ran a series of simulated agent-instructed assembly tasks in which participants were verbally instructed to construct two different Lego models. These were conducted though a mixed design approach. The first twenty-four participants were tested for the two agent conditions within-subjects (vague and non-vague) and the task condition between-subjects (stress and no-stress). This was partially balanced with the subsequent six participants being given the reverse: the agent conditions were between-subjects and the stress conditions within-subjects.

Each session was filmed from two angles. The native camera on a MacBook Pro 10.2, which provided the interfaces for each task, was set to record the entire session to capture the front facing angle of each task. This allowed for the recording of participant facial gestures. This was also the same machine that provided the interfaces for the assembly tasks. A Panasonic HDC-SD80 camera was also set up to record each session from the side to allow for a more detailed view of the model assembly (see Figure 1).

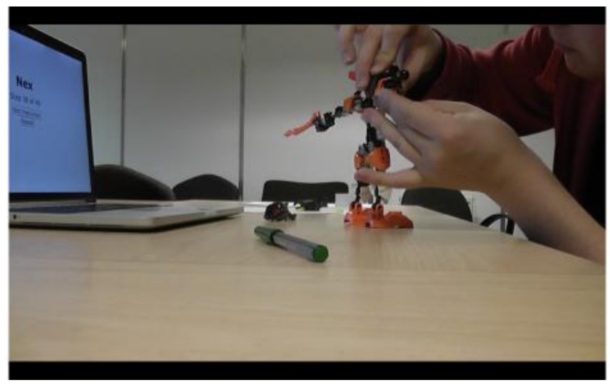

Figure 1. A participant constructs the model *Nex* in front of the agent interface.

Rather than develop an actual agent as such we created simulated agent interfaces that would provide a similar experience for the users. To create these agents each model was first split into 48 steps to produce the non-vague instructions. These were then altered to include VL items outlined in the framework to create the vague agent (Table 1).

	Non-Vague Agent	Vague Agent
Step 17 ***Nex***	Twist this piece so the fin is pointing towards the desk	**Just** give this piece a **little bit** of a twist so the fin is **more or less** pointing towards the desk

Table 1. A comparison of non-vague and vague agent instructions from Step 17 of the model *Nex*.

Each set of instructions was inputted into the Text2SpeechPro software (http://www.hewbo.com) using the synthesised voice Cepstral Lawrence (http://www.cepstral.com) and exported as individual .mp3 files. Four separate HTML files were then created for the two versions of each model. These files functioned as the simulated agent interfaces and gave participants the options to request the next instruction and simulated agent interface and gave participants the options to request the next instruction and repeat the current one (Figure 2). As well as the interface they also functioned as tools for logging the number of repetitions requested in each task and the time taken to complete them.

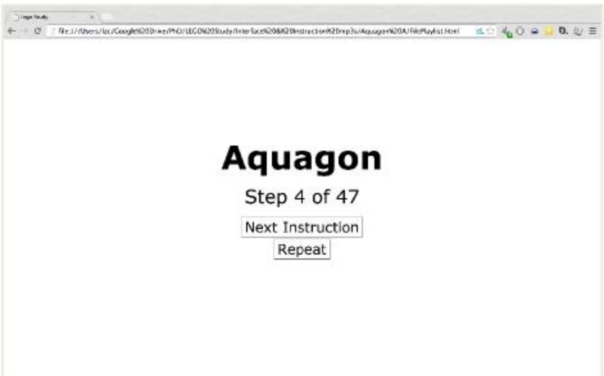

Figure 2. The user interface for the model *Aquagon.*

For designing the two task conditions in the first 12 participants were given the no-stress condition and were timed in all three of their tasks, including the practice model. The mean average of these times was then calculated and each one increased by two minutes and thirty seconds. This total was used as the time limits for participants in the stress condition though they were not informed of the specific time, only that a limit was in place. As the research questions were primarily concerned with exposing subjects to vague language this time increase allowed for greater confidence in the majority of them finishing the task and being exposed to every instruction equally.

Participants

Thirty native English speakers studying at the University of Nottingham were recruited for this study and reimbursed with a £10 voucher for their participation. Nineteen students were male (63.3%) and eleven were female (36.7%). Of these, five were postgraduates and twenty-five were undergraduates. The ages of the participants ranged from 18-30 years old.

Procedure

Following a briefing participants were first tasked with constructing a practice Lego model using the manufacturer's instructions. Those undergoing the no-stress condition were told they had as much time as necessary for all tasks while those in the stress condition were made aware of the time limit known only to the researcher. The practice was followed by two further models using the agent instructions, with each of these proceeded by a questionnaire and interview.

Measures

Both quantitative and qualitative measures were used to assess the interactions with the agents.

Agent Perception

A five point Likert scale was used in post-task surveys to assess how participants rated the agent across eight attributes modified from an existing voice attribute scale [23] – likeable, friendly, trustworthy, sociable, controlling, authoritative, clear and direct – based on the hypotheses described earlier.

Open-ended questions in both the survey and semi-structured interviews were used to gain a greater understanding of their experience and attain a richer detail to as to their perception of the agent, thoughts on the language it used and if and how they would consider interacting with it again. An iterative content analysis approach was used to develop themes from this data.

Task Performance

Performance consisted of two measurements – the time taken to complete the task and the number of steps repeated.

RESULTS

Survey Measures

The results show that H2 was only partially correct. A mixed-design ANOVA revealed the non-vague agent was rated as more direct than the vague, but this was not affected by stress (Figure 3), $F\ (1, 26) = 14.62$, $p < .001$, $\eta_p^2 = .38$. Similarly, the non-vague agent was rated as more authoritative than the vague agent, which again was not affected by stress, $F\ (1, 26) = 15.79$, $p < .001$, $\eta_p^2 = .36$. The other hypotheses were not observed; no other attributes had a significant difference nor did any occur when comparing across the stress and no-stress conditions.

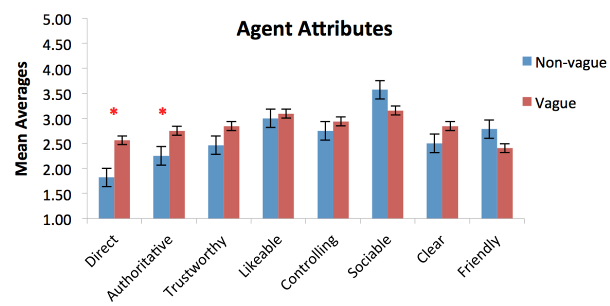

Figure 3. Comparison of vague and non-vague agent attributes measured in the survey showing the mean averages (*1= strongly agree, 5= strongly disagree).* Significant results are denoted with asterisks

User Performance

There was no marked difference in user performance across the two agent conditions (H3), however a one-way ANOVA revealed participants were seen to request less repeats in the stress condition than those in the no-stress $F\ (1, 59) = 5.97$, $p < .05$. There was no significant difference in time taken to complete the task.

Emerging Themes

These themes arose from the combination of the qualitative data discussed in the previous section. We observed no difference between the stress and no-stress conditions and as such this section focuses on the differences and similarities we observed between the agents.

Lexis

Almost all participants who interacted with both agents were able to identify lexical differences between them. The most commonly identified vague items were *basically*, *like* and *just*. These were followed by *kind* of, *sort of* and *more or less*, though participants mentioned *should*, *so* or *now*. For the direct tasks there was no explicit mention of any language.

Responses to particular words were mixed. The use of *basically* in particular appeared to cause an unfavourable perception of the agent, with it described as "inappropriate and somewhat demeaning" as well as being "annoying and too vague [and creating] a condescending tone". Similar responses were seen with *just*: "If I had this in my sat nav I would probably crash my car". Unlike *basically*, however, it also received positive feedback on its suitability with instructions that matched the procedure: "It was more consistent with the step "like it *was* just a little twist" and in creating a better impression of the agent: "I thought it was friendlier: "it sounded like a natural thing for him to be saying".

Similar mixed responses occurred when referring the to language as a whole, again creating a negative impression of the vague agent: "It felt like it was insincere"; "It seemed fake: "it was trying too hard", indicating a certain lack of success in creating a social leveling effect.

When participants were able to compare both agents there was a strong preference for the direct alternative due to its lack of vague language: "I liked the lack of fluffy words". Its simplicity in language was also praised: "It just said what was needed"; "It was much more straightforward so you just do what it tells you", indicating participants sometimes found the VL superfluous to the instructions as a whole.

VL Frequency

Although there was a mixed response to the vague agent, there were indications that reducing the frequency of VL in its instructions could change user perception of the agents: "I wouldn't mind it if they didn't say *like* so much"; "It's okay but think it says just too much"; "It's used too heavily", also pointing to a lack of variety in the language.

Voice Quality

There was a notable difference in how participants responded to the voice of the agent when the language changed. The non-vague tasks received better appraisals: "It was fine. I wouldn't really change anything"; "It's kind of what you expect from a computerised voice". Pairing the same voice with VL however produced difficulty in accepting the agent: "It's just the combination of the voice and script didn't work"; "It sounded too forced".

There was a general consensus that improving the voice would directly impact how the VL was received: "It would sound better with a more natural voice"; "The voice is holding it back". Specific recommendations included the addition of more sophisticated prosodic features: "Change the speed since it's quicker in human speech"; "It emphasized phrases like *more or less* strangely". This indicates a degree of technical limitation being on obstacle in the interactional success of the vague agent.

Context

The instructive nature of the task did not always combine well with the use of VL. It appeared to interfere with task performance and in turn agent perception: "I just wanted to get the job done"; "The extra information was meant to help but it ended up being confusing". Conversely, the non-vague agent received praise on the appropriateness of its instructions: "I think direct is just the way it has to be when getting instructions".

Using a vague agent in other contexts outside instructions received positive feedback, given it attends to the needs of the task at hand: "It's fine so long as it doesn't impact on what needs to be done"; "If it was for something that wasn't so precise [as instructions] then it'd work".

Multi-Modality

Having speech as almost the entire agent interaction had its drawbacks. The practice model using visual instructions resulted in the spoken instructions being less well received in some aspects of the task: "They're [visual] easier to relate to"; "Visual is easier for locating the right piece". Despite this, some elements were made easier with the spoken agents: "Verbal was easier for the actual assembly"; "It was easier to navigate around the 3D space with the spoken". This led to calls in combining the two to reap the positive aspects of each medium: "A visual supplement would make things easier"; "A mixture of both would be nice".

DISCUSSION

The aim of this study was to investigate the suitability of an agent instructing humans using vague language and its predicted effects on agent perception and task performance, using previous literature to create the language framework.

Responses from participants regarding the attributes of both agents were less definitive than in research studying advice-giving interactions, with only the key differences found only in the *direct* and *authoritative* characteristics. These two were expected as the lack of VL in the non-vague agent naturally creates a direct tone. This in turn produces an air of authority and a gap in social power between speaker and receiver. User performance did not vary significantly across the two agent conditions, though less repeats were used in the stress condition. This is likely a result of the unknown time limit creating a sense of urgency, leaving less time to check instructions again and perhaps forcing participants to employ greater focus during those tasks.

Though the results did not come out as expected, the interview data yielded a rich insight into why this may have occurred, as well as generating a greater understanding of the interaction as a whole. Some VL items, for example, were better received than others, either due to their appropriateness or frequency, but this was not overly consistent. VL was less well received in instructions than has been shown in advice giving, perhaps due to the inflexible nature of the outcomes that may warrant a more direct approach. The findings suggest, however, that there is also significant amount of individual preference at play. Some participants praised the use of VL, with several responses mimicking those seen in human interactions. Language is not a static entity that always warrants a general approach and perhaps more variety in the degrees of precision and imprecision are required. We must also strongly consider the individual user and their own preferences in agent design.

Other obstacles need to be overcome if VL is to be used successfully. The technical capabilities of the agent's voice were a strong barrier in acceptance as opposed to just the language. While VL is commonly used feature in everyday language, there is very little exposure to its use by an agent. As the agent lacked the vocal characteristics such as intonation and stress that accompany language, combining the two to create something more familiar could have greater success than either of them individually.

Limitations and Future Work

This study only used a simulated agent to create an interaction. It is unknown whether one agency could the change the outcomes seen here. Similarly, there could be significant differences with introduction of features such as embodiment and alternative exposure times (particularly in relational agents that may see interaction over a period of weeks). Investigating the multiple permutations available in these interactions would provide us with greater insight into how we may benefit from a greater awareness of agent language use.

Despite the mixed responses in this context, the data reveals there is potential in HAI for a greater human like approach to language. This study points towards the development of adaptive agents in the future – those that change their language depending on context and the individual user. For now though, further research in both instructive and non-instructive contexts is required so we may fully begin to understand the potential of alternative communicative strategies.

CONCLUSION

This study compared user reactions to vague and non-vague agent instructors. We ran a series of simulated agent-instructed assembly tasks to discover how agent perception is affected and how the interaction is experienced differently. Findings suggest that there is no one size fits all approach to language in agent design and there must be awareness of the context of the interaction, technical capabilities of the agent and the preferences of the individual user. Given the mixed responses there is potential for vague language in human-agent interaction, but these obstacles must first be researched in greater detail to achieve success.

ACKNOWLEDGEMENTS

This work was supported by EPSRC Grant No. EP/I011587/1.

REFERENCES

1. Adolphs, S., Atkins, S. and Harvey, K. Caught between professional requirements and interpersonal needs: Vague language in healthcare contexts, *Vague Language Explored* (2007), pp. 62-78.
2. Ball, M. and Callaghan, V. Introducing Intelligent Environments, Agents and Autonomy to Users. *IEEE* (2011), pp. 382-385.
3. Bickmore, T. and Cassell, J. Relational Agents: A Modl and Implementation of Building User Trust. Proc. *Human Factors in Computing Systems*, ACM Press, New York, NY (2001), pp. 396-403.
4. Bickmore, T. and Cassell, J. *Social Dialogue with Embodied Conversational Agents, Advances in natural multimodal dialogue systems*. Springer (2005), pp.23-54.
5. Brown, P. and Levinson, S.C. *Politeness: Some universals on language usage*. Cambridge University Press (1987).
6. Bulitko, V.V. and Wilkins, D.C. Automated instructor assistant for ship damage control. *AAAI/IAAI* (1999), pp.778-785.
7. Carter, R. Orders of reality: CANCODE, communication and culture. *ELT Journal,* 52 (1998), pp. 43-56.
8. Chan, V., Ray, P. and Parameswaran, N. Mobile e-Health monitoring: an agent-based approach, *IET communications*, 2 (2008), pp.223-230.
9. Channel, J. *Vague Language*. Oxford University Press (1994).
10. Cutting, J. *Vague Language Explored*. Palgrave Macmillan (2007).
11. Doswell, J. and Harmeyer, K. Extending the serious game boundary: Virtual instructors in mobile mixed reality learning games. *Digital Games Research Association International Conference*, Citeseer (2007)
12. Goffman, E. *The presentation of self in everyday* life. Garden City, NY (2002).
13. Goffman, E. *Interaction Ritual: Essays on Face-to-Face Behaviour*. Anchor Books (1967).
14. Heylen, D., Nijholt, R., Den Akker, O.P. and Vissers, M. Socially intelligent tutor agents. *Intelligent Virtual* Agents, Springer (2003), pp.341-347.
15. Koester, A. 'About twelve thousand or so': Vagueness in North American and UK offices, *Vague Language Explored* (2007), pp. 40-61.

16. Lakoff, G. *Hedges: A Study in meaning criteria and the logic of fuzzy* concepts. Springer (1975)

17. Maes, P. Agents that reduce work and information overload, *Communications of the* ACM, 37 (1994), pp.30-40.

18. McCarthy, M. and Carter, R. *As visible patterns of interaction*, Explorations in Corpus Linguistics (2006).

19. Prince, E.F., Frader. J. and Bosk, C. On hedging in physician-physician discourse. *Linguistics and the* Professions (1982), pp.83-97.

20. Rowland, T. 'Well maybe not exactly, but it's around fifty basically?' Vague language in mathematics classrooms. *Vague Language Explored* (2007), pp.79-96.

21. Schaafstal, A. M., Johnston, J. H. and Oser, R. L. Training teams for emergency management, *Computers in Human* Behaviour, 17 (2001), pp.615-626.

22. Sukthankar, G., Shumaker, R. and Lewis, M. Intelligent agents as teammates, *Theories of Team Cognition: Cross-Disciplinary Perspectives* (2012), pp.313-343.

23. Tamagawa, R., Watson, C.I., Kuo, I.H., Macdonald, B.A. and Broadbent, E. The Effects of Synthesized Voice Accents on User Perceptions of Robots, *International Journal of Social* Robotics, 3 (2011), pp. 253-262.

24. Torrey, C., Fussell, S. and Kiesler, S. How a robot should give advice, *Proceedings of the 8th ACM/IEEE international conference on Human-robot* interaction, IEEE Press, Tokyo, Japan (2013), pp. 275-282.

25. Trappes-Lomax, H. Vague language as a means of self-protective avoidance: Tension management in conference talks. *Vague Language Explored* (2007), pp.117-137.

26. Wang, Wang, N., Johnson, W.L., Mayer, R.E., Rizzo, P. Shaw, E. and Collins, H. The politeness effect: Pedagogical agents and learning outcomes, *International Journal of Human-Computer Studies,* 66 (2008), pp. 98-112.